***BUS 347.01 Introduction to Business Analytics***

***Homework 3 (60 Points)***

**Homework Description:**

The attached file “HW3 Data.csv” contains data records regarding the “pick errors” (i.e., retrieving the wrong item) in the distribution center for an online retailer. The warehouse manager thinks most errors are due to inexperienced workers and she believes that a training program also may help to reduce them. Before sending all employees to training, she examines data from a pilot study of 30 employees (half attended training and the other half did not) to decide whether the training program will be effective.

You need to perform a regression analysis based on this dataset to the elements that can potential reduce the “pick errors” through answering the following questions. Your submission should be an analytics report in the word format knitted from RMarkdown code.

In your analytics report, you need clearly label the following items:

* Question Number
* R Code
* R Output
* Conclusion, if applicable

You need to submit the knitted analytics report on Blackboard. Please carefully check your work before the submission, as you can only submit your work once. Late submissions will not be accepted.

**Homework Questions (Each Question has 10 point)**

For each question, use the whole sample as training dataset.

**Q1.** Fit a simple linear regression . Report the R-Squared, Adjusted R-Squared and RMSE for Model 1.

**Q2.** Fit a multiple linear regression . Report the R-Squared, Adjusted R-Squared and RMSE for Model 2.

**Q3.** Fit a second multiple linear regression . Report the R-Squared, Adjusted R-Squared and RMSE for Model 3.

**Q4.** Fit a third multiple linear regression . Report the R-Squared, Adjusted R-Squared and RMSE for Model 4.

**Q5**. Based on the previous model fits, discuss i) which model provides the best model structure and why; ii) which model provides the best forecasting accuracy and why.

**Q6.** Based on the model that provides best forecasting accuracy, what will be the predicted pick errors for the following four people?

* A: Yrs.Exp=10, Training=1
* B: Yrs.Exp=20, Training=0
* C: Yrs.Exp=5, Training=1
* D: Yrs.Exp=1, Training=1